REDUCING UNCERTAINTY IN SITE CHARACTERIZATION USING BAYES MONTE CARLO METHODS

By Michael D. Sohn,[1] Associate Member, ASCE, Mitchell J. Small,[2] Member, ASCE, and Marina Pantazidou,[3] Associate Member, ASCE

ABSTRACT: A Bayesian uncertainty analysis approach is developed as a tool for assessing and reducing uncertainty in ground-water flow and chemical transport predictions. The method is illustrated for a site contaminated with chlorinated hydrocarbons. Uncertainty in source characterization, in chemical transport parameters, and in the assumed hydrogeologic structure was evaluated using engineering judgment and updated using observed field data. The updating approach using observed hydraulic head data was able to differentiate between reasonable and unreasonable hydraulic conductivity fields but could not differentiate between alternative conceptual models for the geological structure of the subsurface at the site. Updating using observed chemical concentration data reduced the uncertainty in most parameters and reduced uncertainty in alternative conceptual models describing the geological structure at the site, source locations, and the chemicals released at these sources. Thirty-year transport projections for no-action and source containment scenarios demonstrate a typical application of the methods.

INTRODUCTION

This paper presents a probabilistic alternative to the more traditional calibration process of deterministic models used for analyzing flow and transport at hazardous waste sites. In a deterministic numerical analysis, the conceptual flow and transport model is first specified, and then the model parameters are fine-tuned through calibration, a nonunique process based on the comparison of selected model results and observations. The Bayes Monte Carlo (BMC) approach presented herein is a structured method of using existing data to reduce the uncertainty in both the conceptual models and the model parameters. Similar to a typical calibration process, some decisions involved in the application of the proposed BMC method are not unique.

The BMC approach consists of two steps. First, suitable conceptual models and the distributions of model parameters are identified. This selection process is based on the best estimate that the model developer can provide prior to comparing model results to actual observations. For each conceptual model, multiple model realizations are generated at this stage, by sampling the space of the model parameters following Monte Carlo or other sampling techniques. The ''prior'' refers to the distribution of the parameters and the model results at this step. The a priori uncertainty of the problem is described by the prior distribution of the model parameters and model results, or suitable summary statistics, such as the parameter standard deviations.

In the second step, the model results are compared with existing observations through a structured probabilistic methodology referred to as Bayesian updating. Although the Bayesian updating methodology is well defined, the way in which observations are compared with model results involves an assessment and decisions that depend on how the actual field data are related to model predictions. These judgments account for the fact that, by necessity, any model involves simplification of real-world processes and some degree of spatial and temporal aggregation. The procedure for comparison and the resulting update are thus not unique. Alternative approaches for this comparison are discussed in detail in the paper.

Once the Bayesian methodology is applied, an updated, or posterior, probability is computed for each realization, given the set of observations. At the end of the second step, these calculated probabilities provide weights for both the model parameters and the model results, combining the information contained in the observed data with that of the prior estimates. All reweighting occurs on the original set of simulations, so no additional model runs are required at this stage; as a result, additional observations can easily be incorporated in subsequent updating. The posterior distributions of model parameters and model results reflect the updated uncertainty.

The objectives of this paper are to provide details on the development of the Bayesian approach and the decisions involved in applying the methodology for ground-water site characterization. Using data from an actual hazardous waste site, this paper demonstrates that a BMC methodology provides an effective tool for characterizing uncertainty, and for incorporating new information for the purpose of reducing this uncertainty.

BACKGROUND

Many existing studies have addressed the problem of uncertainty in modeling subsurface flow and transport phenomena using inverse modeling (e.g., Yeh 1986) and optimization (e.g., Marryott 1996). While significant, these studies focus primarily on reducing uncertainty in model parameters and do not address the combined uncertainty in conceptual models and model parameters simultaneously. Monte Carlo techniques (e.g., Essaid and Hess 1993) provide a natural method for characterizing uncertainty in both conceptual models and model parameters, but they do not provide a formal method for comparing model predictions with field data. In addition, when additional information on model parameters or field measurements are gathered, new parameters must be reestimated and model simulations recomputed. This can result in a significant computational burden if ongoing data collection occurs at a study site.

Bayesian statistical methods have been used to address these difficulties in a variety of environmental applications (see Sohn 1998 for a comprehensive review). In ground-water ap-

[1]Postdoctoral Fellow, Envir. Energy Technol. Div., Lawrence Berkeley National Lab., One Cyclotron Rd., M/S 90-3058, Berkeley, CA 94720; formerly, Grad. Student, Dept. of Civ. and Envir. Engrg., Carnegie Mellon Univ., Pittsburgh, PA 15213. E-mail: mdsohn@lbl.gov

[2]Prof., Dept. of Engrg. and Public Policy, Carnegie Mellon Univ., Pittsburgh, PA.

[3]Asst. Prof., Dept. of Civ. and Envir. Engrg., Carnegie Mellon Univ., Pittsburgh, PA.

Note. Associate Editor: Susan E. Powers. Discussion open until March 1, 2001. To extend the closing date one month, a written request must be filed with the ASCE Manager of Journals. The manuscript for this paper was submitted for review and possible publication on April 22, 1999. This paper is part of the ***Journal of Environmental Engineering***, Vol. 126, No. 10, October, 2000. ©ASCE, ISSN 0733-9372/00/0010-0893–0902/$8.00 + $.50 per page. Paper No. 20783.

plications, Bayesian methods have been used to incorporate expert judgment in identifying and estimating initial uncertainty estimates and the uncertainty reduction achievable following comparisons with observed field data (e.g., Loaciga and Mariño 1987; Reichard and Evans 1989; Massmann et al. 1991; James and Freeze 1993). While these studies have advanced the state of the art for uncertainty characterization and reduction, their primary focus has been on the stochastic variability and uncertainty associated with subsurface heterogeneity in hydraulic conductivity. They have not considered in much detail the characterization of other significant chemical and biological processes that control the dissolution, fate, and transport of ground-water contaminants. As such, they have not investigated the full range of geochemical and conceptual model issues related to uncertainty in site characterization and remedial system design. This work further develops the Bayes Monte Carlo (BMC) method to update uncertainty, not only in the hydrogeologic parameters describing ground-water flow but also in the assumed geological structure of the site, the location and characterization of the primary chemical sources, and the chemical fate and transport parameters.

BAYES MONTE CARLO UPDATING

The traditional type of uncertainty analysis, where uncertainties in model input parameters are routed through a predictive model using analytical or numerical (e.g., Monte Carlo) methods, yields only a prior, or initial, assessment of uncertainty. The BMC approach allows the prior to be updated to a posterior probability using Bayes' rule to account for the level of agreement between model predictions and the observed field data. The posterior probability of each realization of the Monte Carlo simulation, given the observed data, is computed as (Brand and Small 1995)

$$p'_k = \frac{L(O|Y_k)p(Y_k)}{\sum_{i=1}^{U} L(O|Y_i)p(Y_i)} \quad (1)$$

where p'_k = posterior probability of the kth Monte Carlo simulation; $L(O|Y_k)$ = likelihood of observation O given the model prediction Y_k; $p(Y_k)$ = prior probability of the kth Monte Carlo simulation; and U = number of the Monte Carlo simulations. If many independent observations are considered, then $L(O|Y_k)$ in (1) represents the likelihood of observing all of the observations simultaneously, i.e.

$$L(O|Y_k) = \prod_{s=1}^{S} L(O_s|Y_k) \quad (2)$$

where S = number of independent observations. The likelihood function quantifies the difference between the observations and the model output resulting from inherent variability, measurement error, spatial and temporal averaging, and imperfect model representation. For example, for unbiased measurements with a normally distributed error, the likelihood of an observation is given as

$$L(O|Y_k) = f(O - Y_k) = \frac{1}{\sqrt{2\pi}\sigma_\varepsilon} \exp\left(-\frac{1}{2}\left[\frac{O - Y_k}{\sigma_\varepsilon}\right]^2\right) \quad (3)$$

where σ_ε^2 = observation error variance. The posterior probability, p'_k, of each realization is applied to the full set of model assumptions and predictions associated with that realization, including the input parameter vector (e.g., physical and chemical model input parameters) and the conceptual model used for that realization. Hence, this posterior probability can be used to reestimate the uncertainty distribution for each component of the model. The posterior mean, variance, and correlation coefficient of each model output and input parameter are estimated from the posterior probability of each realization as (Brand and Small 1995)

$$\mu'_V = \sum_{i=1}^{U} V_i \cdot p'_i \quad (4)$$

$$\sigma'^2_V = \sum_{i=1}^{U} (V_i - \mu'_V)^2 \cdot p'_i \quad (5)$$

$$\rho'_{V,W} = \frac{\sum_{i=1}^{U} (V_i - \mu'_V)(W_i - \mu'_W) \cdot p'_i}{\sigma'_V \cdot \sigma'_W} \quad (6)$$

where V and W represent any model input or output. The posterior correlation coefficient, $\rho'_{V,W}$, provides insight into the parameter correlations and the combinations that can yield similar model output—e.g., parameters that affect the model predictions in a similar manner tend to a negative correlation induced, while parameters that have offsetting effects tend to have positive correlation.

As ongoing information is obtained (e.g., following regular field sampling events), the computed posterior from one round of updating becomes the prior for the next, i.e., the $p(Y_i)$ in the subsequent update is set equal to the p'_k from the previous update. So long as the original sampling provides full and adequate coverage of the model and input parameter sample space, the updating can be conducted without further evaluation of the flow and transport models.

Method-Specific Issues

The selection of the appropriate error structure for the likelihood function is a key consideration for the BMC method; it requires a careful consideration of the relationship between the model predictions and the observed data. For example, when the relationship is direct—i.e., when model predictions and observed data are available at the same level of temporal and spatial aggregation—then a likelihood estimate based on field and laboratory measurement error is appropriate. In many applications, the correspondence between observed data and model predictions is less direct. For example, steady-state flow models predict constant ground-water head and flow for conditions of constant infiltration, recharge, and pumping. Similarly, discretized models for contaminant transport yield representative concentrations at the spatial scale of the grid, ignoring finer-scale variations in environmental properties and parameters, and the resulting finer-scale variations in concentration. In these cases, the error variance in (3) must incorporate the effects of the unrepresented spatial and temporal variations as well as the error associated with inaccuracy in the measurement methods used to obtain the data. These issues are exemplified by considering some of the decisions made in the application presented in the second part of this paper: (1) Coarse-scale flow models are used yielding predictions of hydraulic head and chemical concentration at points centered on a 50 m grid spacing; (2) a single average value is computed for all the existing head data within each grid cell, assuming steady-state ground-water flow; and (3) chemical concentration measurements and predictions are averaged within one-year intervals. It is probable, however, that the significant variations (especially for concentration) that are observed within a year and at smaller spatial scale than that of the model grid are in fact the dominant component of the error variance. In this application, multiple samples within a year and within the same cell indicate an error structure for the observed head that is normally distributed, and an error structure for measured chemical concentrations that is lognormal. For the latter, O

and Y in (3) correspond to the logarithms of the observed and modeled values, respectively, and σ_ε^2 is the variance of the error of the logarithms. Methods for estimating σ_ε^2 are described in the application.

An additional complication arises from the independence assumption, which is inherent in the use of (2). This assumption, though commonly employed, is often inappropriate. Spatial and temporal correlations in the observed data, model predictions, and the differences between them, can violate the independence assumption, causing the information content of an observed dataset to be reduced. Classical methods for testing the hypothesis that the observed data are sampled from a ''world in which the proposed model is true'' account for this by reducing the effective sample size of the test statistic (see, e.g., Yevjevich 1972; Chapra and Reckhow 1983). In our application, a similar approach is employed: the likelihood is defined using statistics that represent the aggregate relationship between the observed chemical data and the model, and the effective sample size is adjusted to account for the spatial correlation of the measured data.

Application-Specific Issues

Applications of the proposed methodology require decisions specific to the phenomena studied and the available data. The decisions involved in this ground-water flow and chemical transport application offer to researchers and practicing engineers several alternatives in implementing the methodology. The large site studied in this application, which includes several sources and is characterized by a considerable quantity of varied field measurements, was discretized at a course scale (a 50 m grid). Each of the adjustments required to predict the regional flow and transport processes using the small-scale measurements are described in detail at each updating sequence. Noteworthy decisions include (1) conducting a two-step update, the first for the head data and the second for the chemical data; (2) using a ''Bayes window'' approach in the head updating because the direct likelihood estimates given by Eqs. (2) and (3) appeared to overly discriminate between several well-fitting realizations; and (3) introducing an alternative likelihood estimate in the chemical updating to account for data correlations. The Bayes window and the alternative likelihood estimate both provide useful tools that should be evaluated on a case-by-case basis when implementing the BMC methodology.

APPLICATION TO A CHLORINATED SOLVENT PLUME

Site Description

The BMC method is demonstrated for a site with multiple sources of chlorinated hydrocarbons in ground water. The study area (Fig. 1) includes four Superfund sites and is comprised of several industrial and federal facilities with chemical sources contributing to a chlorinated solvent plume that extends over an area of approximately 3.7×0.9 km. Notable features include many potential contaminant sources, which can be grouped in three main source areas (SA-1, SA-2, and SA-3), extensive paved areas that limit sampling (a portion of a freeway and two aircraft runways), and a regional groundwater flow gradient toward a nearby bay and wetlands. The upper two of the four major permeable zones identified in boring logs contain the vast majority of the contaminant mass. The first layer is unconfined and the second layer is confined. The available head data suggest significant hydraulic interaction between these two aquifer zones.

Chemicals in SA-1 and SA-3 are not believed to have been released until approximately 1960, while SA-2 sources are believed to have been active since approximately 1940. In 1987,

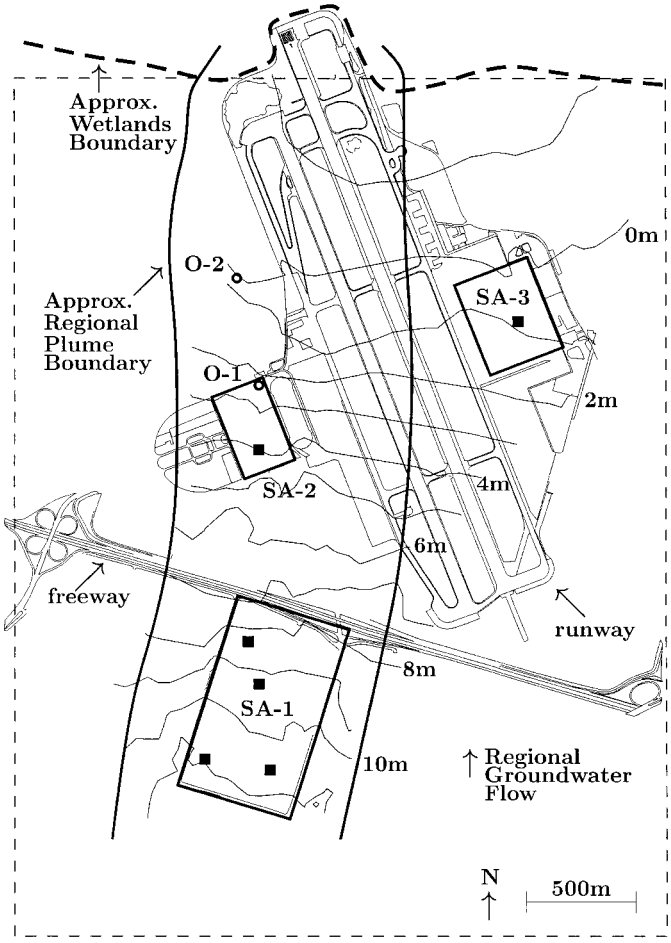

FIG. 1. Illustration of Study Site. Dotted Box Outlines Model Area; SA-1, SA-2, and SA-3 Are Descriptive Labels for Source Areas, with Black Squares Identifying Areas of Model Source Cells; O-1 and O-2 Are Cells Where Model Predictions Are Plotted and Compared with Observed Data; Contoured Lines Indicate Observed Mean Head in Uppermost Permeable Zone

slurry walls were installed in SA-1 to prevent further migration of contaminants from the source areas in combination with localized pumping. The significant chemicals of concern at the site include chlorinated hydrocarbons, in particular, trichloroethylene (TCE) and 1,2-dichloroethylene (DCE). The presence of dense, nonaqueous phase liquid (DNAPL) residuals is suspected at various locations, which can be assumed to serve as sources for the site plume. While TCE is known to have been discharged from various site activities, it is not known whether DCE was discharged at the source locations or is present in the aquifer solely as a decay product of TCE degradation. Discrimination between alternative source scenarios (i.e., source type and source location) is one of the objectives of our analysis.

Overview of Modeling Procedure

The contaminant distribution in the two upper aquifers is investigated using the ground-water flow model MODFLOW (McDonald and Harbaugh 1988), coupled with the contaminant transport and fate model MT3D (Zheng 1992); both were processed on UNIX-based workstations. The workstation processing speed was approximately one-fourth the speed of current personal computers so should not pose a significant computational burden; our simulations required approximately one week on four workstations.

A two-stage updating procedure was used to simulate the uncertainty in the ground-water flow and contaminant trans-

port. In the first stage, a prior characterization of the spatial distribution of hydraulic conductivity was determined from well-boring logs. An ensemble of equally likely hydraulic conductivity fields was generated and subsequently updated using head measurements. The second stage involved defining a set of chemical transport parameters to represent the uncertainty in source, dispersion, sorption, and reaction terms. Chemical transport predictions were generated and subsequently updated using chemical concentration measurements. The posterior probabilities were summed across corresponding hydraulic conductivity, source scenario, and chemical parameter sets to determine the relative posterior support for the alternative model assumptions and parameter values. These posterior estimates provide a basis for characterizing uncertainty in both current conditions at the site and predicted future concentrations.

Flow Model

The upper aquifer at the site was simulated as an unconfined aquifer extending from approximately 10 m above to 5 m below mean sea level (msl). The second aquifer was modeled as a confined aquifer to a depth of 17 m below msl. The simulation domain was discretized on a uniform 50 m square grid, yielding 9,600 grid cells. Hydraulic conductivity estimates for each grid cell in each of the two aquifer layers were derived from 700 boring logs in and around the study area. Both directionally isotropic and anisotropic semivariograms were used to characterize the spatial correlation structure of the hydraulic conductivity. There is a higher degree of spatial correlation in the longitudinal (north-south) direction than in the transverse (east-west) direction, suggesting an anisotropic correlation structure. Nonetheless, to consider alternative conceptual representations of the subsurface geology and correlation structure, both isotropic and anisotropic hydraulic conductivity fields were simulated for the site. Sequential Gaussian simulation was used to generate 1,000 realizations for each of the isotropic and anisotropic hydraulic conductivity structures. The realizations were conditioned on the respective semivariograms and the observed hydraulic conductivity estimates using the GSLIB software package (Deutsch and Journel 1992).

Updating Using Head Data

Fig. 1 illustrates the approximate steady-state head contours from head measurements in the upper aquifer (Layer 1). Similar values are observed in the second layer (Layer 2), as expected from their significant hydraulic connection. The decisions involved in this application of the BMC methodology include the following:

- The head updating was not conducted based on a comparison between measured and predicted head values. Instead, head gradients were estimated using pairs of the head measurements and compared with those estimated from model head predictions between the same locations. This provides (1) a better aggregate assessment of the regional flow, and (2) a more direct relationship to the information used in the chemical transport predictions, since it is the head gradients that are used to estimate the Darcy velocities for the chemical transport, and not individual head values.
- If the difference between the model-predicted and the data-estimated gradients was large, a minimum likelihood value was assigned corresponding to the 95% range of the Gaussian distribution rather than the estimated likelihood value. As a result, any individual gradient comparison resulting in a fit with less than 5% probability was assigned a 5% probability. This decision was made to avoid having an extreme error in one likelihood estimate overshadow the impact of the rest of the observations. This decision can be justified by considering that these extreme errors are more likely to result from a significant measurement error than an error (of such magnitude) in the model prediction.
- A ''Bayes window'' updating approach, described by Franks et al. (1998) in their generalized likelihood uncertainty estimation (GLUE) method, was used to identify the most significant/likely realizations. In a window approach, only the realizations that have significant posterior probabilities are kept and subsequently reassigned equal posterior probabilities, while the others are discarded. In our application, hydraulic conductivity fields with posterior probabilities greater than 0.1% each fit the head data well, and were retained for subsequent analysis. This approach is effective in cases where the updated probabilities may overly discriminate between realizations that each correspond comparably well with the measurements, because of data correlations or spatial aggregation not fully considered in the likelihood function. In the second-stage updating using chemical data, an alternative likelihood estimator is introduced.

The well pairs for gradient calculations were selected so as to maximize the distance between the two wells both longitudinally (north-south) and laterally (east-west), so that regional gradient estimates were used in the updating. Each data point was used once, yielding 165 gradient values in the first aquifer and 66 values in the second aquifer. Likelihood estimates for the updating were conducted using (1)–(6) with the head variances estimated as the pooled error variance (see, e.g., Weisberg 1985). That is, σ_ε in (3) is computed as the standard deviation of the head measurement in that cell for each gradient pair. This error variance characterizes the error from error in the measurements as well as the error associated with comparing cell-averaged model predictions with field measurements (or their statistics) from discrete points within each cell.

Eight simulations, five isotropic and three anisotropic hydraulic conductivity realizations, encompassed 99.9% of the posterior probability. While 95% of the posterior probability was associated with only three of these realizations, the 99.9% cutoff was chosen to include a larger range of variability in the possible hydraulic conductivity fields. All eight of the realizations appeared to fit the data well and were used for the second stage of the updating procedure. Fig. 2 displays the spatial distribution of hydraulic conductivity and the corresponding head predictions for the upper layer for two of the eight ''second-stage'' fields: one isotropic field [Fig. 2(a)] and one anisotropic field [Fig. 2(b)]. The accepted isotropic field appears similar in many respects to the accepted anisotropic field, though the latter indicates a distinctly greater degree of channeling (indicated by areas with light coloration, i.e., high hydraulic conductivity) elongated in the north-south direction.

To further assess the selection of hydraulic conductivity fields by the updating procedure, two additional fields were also carried along to the second stage: (1) the mean hydraulic conductivity field from the anisotropic subset of prior fields [Fig. 2(c)]; and (2) a rejected realization (the hydraulic conductivity field with the lowest posterior probability) [Fig. 2(d)]. The mean hydraulic conductivity field is much smoother in its variation than any of the individual realizations since it is on-average correct, but not reflective of the degree of variation expected at the site. All 10 fields are assigned equal prior probability for the subsequent comparisons of the chemical predictions to the chemical data, in accordance with the

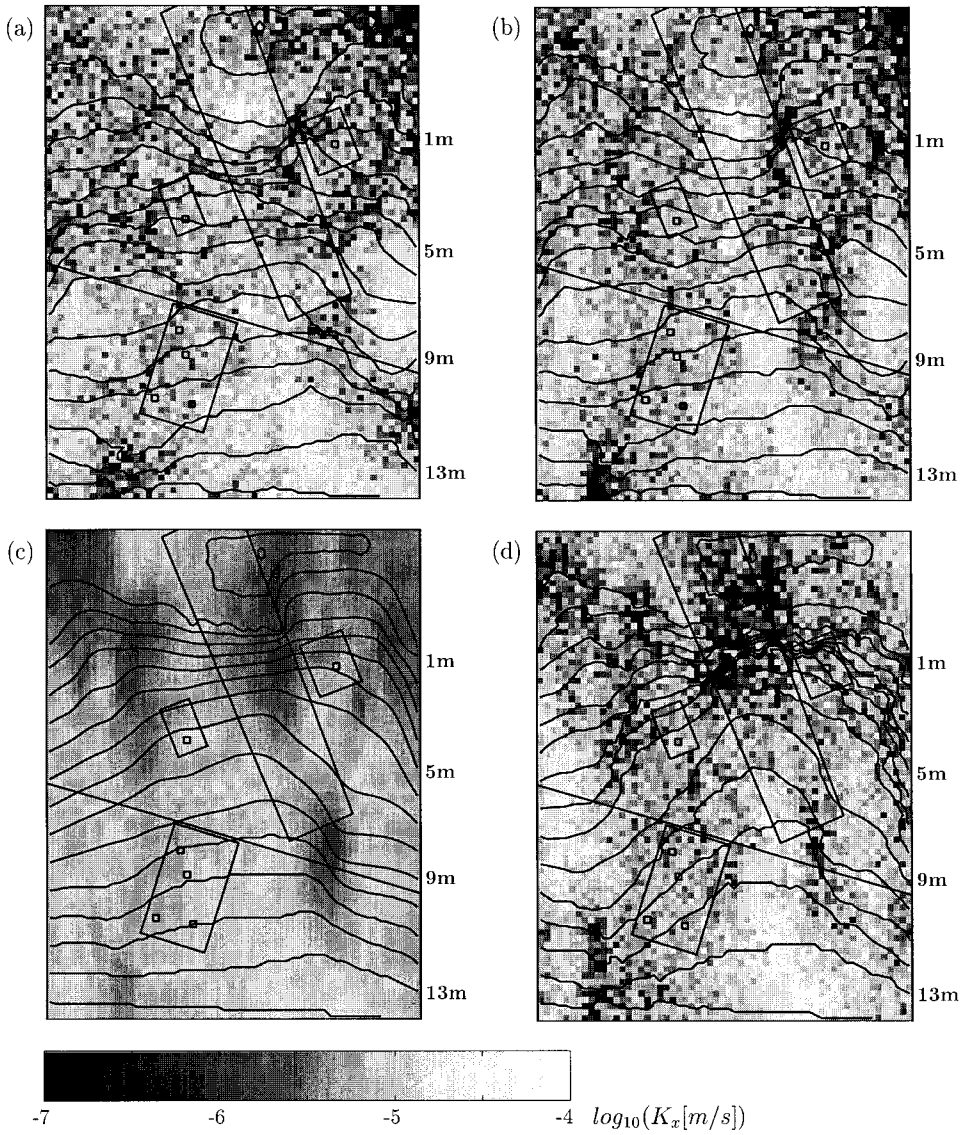

FIG. 2. Four of 10 "Second-Stage" Hydraulic Conductivity Fields and Corresponding Head Predictions: (a) Isotropic Field; (b) Anisotropic Field; (c) Mean of 1,000 Anisotropic Field Priors; (d) Rejected Field

TABLE 1. Summary of Prior Uncertainty Distributions for Chemical Transport Model

Parameter (1)	Range (2)	Distribution (3)
Longitudinal dispersivity (=α_l)	0.1–100 (m)	Uniform
Transverse/longitudinal dispersivity ratio	0.025–1	Log-uniform
Vertical/longitudinal dispersivity ratio	0.003–0.1	Log-uniform
Sorption coefficient (=K_d)	0.16–3.06 l/kg	Log-uniform
Porosity	0.25–0.35	Uniform
Near-source TCE half-life	0.4–50 (years)	Log-uniform
Far-source TCE half-life	0.4–50 (years)	Uniform
Near-source DCE half-life	0.4–50 (years)	Log-uniform
Far-source DCE half-life	0.4–50 (years)	Uniform
Source area concentrations:		
Area 1 TCE initial concentration	50–800 (ppm)	Log-uniform
Area 2 TCE initial concentration	5–40 (ppm)	Log-uniform
Area 3 TCE initial concentration	0.005–0.1 (ppm)	Log-uniform
Area 1 DCE initial concentration[a]	175–4,700 (ppm)	Log-uniform
Area 2 DCE initial concentration[a]	5–40 (ppm)	Log-uniform
Area 3 DCE initial concentration[a]	0.005–0.1 (ppm)	Log-uniform

[a]Applies only to Scenarios 2 and 4, where DCE sources are present (see Table 2).

Bayes window method. While introducing the two additional fields does increase the uncertainty in the hydraulic conductivity fields for the second-stage analysis, it also allows an evaluation of the robustness of the chemical transport updating, as described in the Results section.

Transport Model

The chemical transport model considers uncertainty in the source, dispersion, sorption, and reaction terms. The prior uncertainty distributions for the inputs to MT3D are summarized in Table 1 and briefly described here. A comprehensive description of the model assumptions and parameters is presented in Sohn (1998).

Dispersivity (α)

While the overall scale of the study area is 4 × 3 km, the dispersivity parameters only describe the dispersion that is not accounted for by the intercell heterogeneity in the hydraulic conductivity field. For the model grid spacing of 50 m, the longitudinal dispersivity is estimated to vary between 0.1 and 100 m.

Sorption Coefficient (K_d)

The same value of K_d is assigned to all grid cells and for both TCE and DCE due to a limitation in the MT3D model for sequential reaction modeling. A new MT3D model (Zheng et al. 1997) is currently being developed to evaluate sequential reactions in a direct manner so that this restriction will no longer be necessary. Sorption parameters presented in Howard et al. (1991) and Montgomery (1992), and estimated from localized field experiments conducted at the site (USEPA 1989), indicate that these two chemicals do exhibit similar sorption properties, suggesting that using the same K_d value for TCE and DCE is a reasonable approximation.

Decay

First-order degradation is assumed for both TCE and DCE transformation. All of the TCE degradation is assumed to result in formation of DCE. Separate degradation rates (and uncertainty distributions) for cells near and far away from the sources are used to account for the differences in degradation rates applicable for high and low concentrations, anaerobic and aerobic conditions, sorbed- and dissolved-phase degradation, and possible cometabolic processes in the presence of cocontaminants. It is believed that close to the sources highly anaerobic conditions are possible, which can promote rapid dehalorespiration (e.g., Nyer and Duffin 1997), while further away from the sources, only mildly anaerobic and eventually aerobic conditions can be expected.

In addition to parameter uncertainty, alternative contaminant source scenarios are evaluated in order to determine whether NAPL sources were present in both layers of the aquifer, and whether primary DCE was included in these sources. Four source scenarios, summarized in Table 2, are considered to account for the uncertainty in the location of the sources in the first and second aquifers and in the contaminants released at the source, i.e., either TCE or both TCE and DCE. The four source scenarios are assumed to be a priori equally likely. In SA-1, separate phase TCE is believed to be present. The concentrations at SA-1 were assumed to range between 5 and 75% of the aqueous solubility of TCE and DCE, and are consistent with observed concentrations in the vicinity of the SA-1 sources. Source concentrations for SA-1 are reduced to zero in 1987 to model the source containment activities at the site. For SA-2 and 3, there is little evidence that DNAPLs are present. However, the concentrations in these areas have not significantly changed and little is known about the conditions of the initial contaminant releases. Hence, constant strength source terms are also used, at concentrations varying within a range representative of the measured concentrations at the respective source locations.

Bayes Update Using Concentration Measurements

The overall uncertainty is characterized by the 10 hydraulic conductivity fields from the flow model studies (eight fields from the head data update, the mean field, and a rejected realization), the distributions of chemical transport and reaction parameters in Table 1, and the four possible scenarios for TCE and DCE sources in the upper and lower aquifers, summarized in Table 2. Twelve hundred prior simulations of the MT3D model are used, with 300 assigned to each of the four source scenarios; each hydraulic conductivity field was allocated 30 of these 300. The parameter inputs from Table 1 are sampled using Latin Hypercube sampling. Each of the 1,200 simulations are assigned unique chemical transport and reaction parameters. Application-specific adjustments to the presented BMC methodology include the following:

- Chemical data updating was conducted using data gathered in the year 1986 and 1992. Although historical records include many years (1982–1994), much of the data are obtained at the same locations and are characterized by considerable temporal correlation as well, due to the slow velocity of the plume. The year 1992 was selected because this was the only year in which a comprehensive sampling program was conducted (Fig. 3). The year 1986 was selected because it is sufficiently before the 1992 sampling period to reduce the effects of the temporal correlation, yet contained enough data for a reasonable representation of the regional plume.
- Instead of a comparison between corresponding observations and model predictions, the likelihood function for (1) is defined using two summary statistics—the predicted and observed mean concentrations and the correlation coefficient between the predicted and observed concentrations. This approach is able to capture the essential statistical features of the model-data comparison and to account for the spatial correlation of the measurements.

The proposed likelihood measures are orthogonal, since changes in one can occur without affecting the other; for example, a linear transformation of the model (or the data) will affect the difference in the mean but not the correlation co-

TABLE 2. Summary of Conceptual Model Uncertainty in Location of Sources and Contaminants Released at These Sources (All Scenarios Assigned Equal Numbers of Realizations and Thus Equal Prior Probability in Uncertainty Analysis; Posterior Probabilities Estimated by Updating Prior Probabilities Using Chemical Data)

	Scenario			
Aquifer layer	1	2	3	4
1	TCE	TCE & DCE	TCE	TCE & DCE
2	none	none	TCE	TCE & DCE
Prior probability	25%	25%	25%	25%
Posterior probability	24.9%	14.0%	45.7%	15.4%

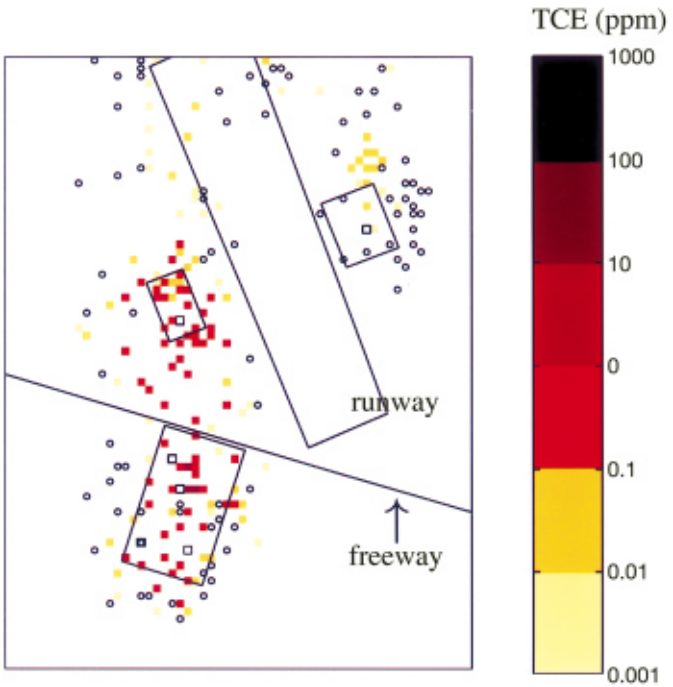

FIG. 3. Average Measured TCE Concentrations within Model Cells for Year 1992 in Layer 1 (Circles Indicate Nondetect for All Measurements in Cell)

efficient. The comparison to the mean ensures that the magnitudes of the predicted plume concentrations are consistent with those of the observed data; the correlation coefficient ensures that the predicted and observed plume shapes are consistent.

The contribution of the mean to the likelihood function is computed from the sampling distribution of the observed mean, given that the mean of the model is true. This sampling distribution is given by

$$\bar{c}_o \sim \text{Normal}(\bar{c}_m, \sigma_\varepsilon^2/N) \tag{7}$$

where $\bar{c}_o$ = mean of the observed natural logarithms (ln) of the concentrations; $\bar{c}_m$ = mean of the model predictions; σ_ε^2 = error variance; and N = sample size.

The sampling distribution of the correlation coefficient, r, between the predicted and observed (ln) concentrations is computed from the transformed variable

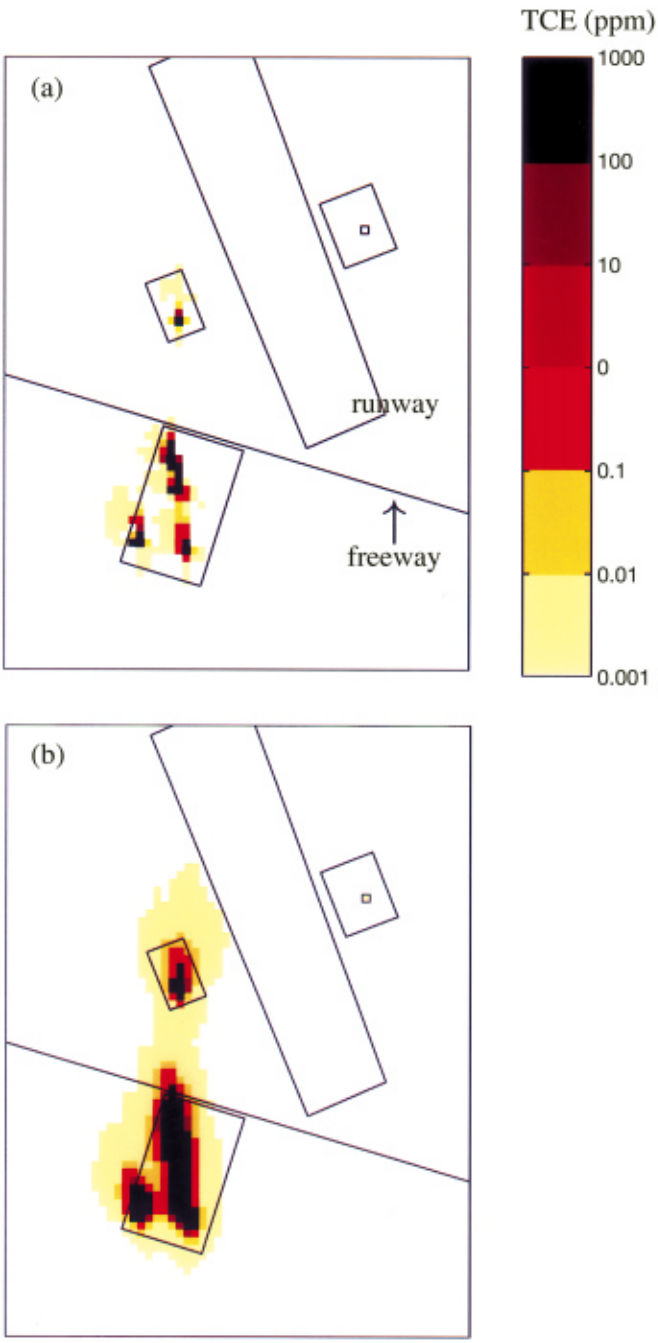

FIG. 4. Prior (a) and Posterior (b) Geometric Mean of TCE Predictions in Layer 1 at Year 1992 from All Simulations

TABLE 3. Summary of Effective Sample Sizes of Available Datasets after Spatial Correlation Is Taken into Account for Chemical Data Update

Dataset (1)	Number of samples (2)	Effective sample size (3)
TCE, Layer 1, 1992	225	20
TCE, Layer 2, 1992	145	11
DCE, Layer 1, 1992	131	5
DCE, Layer 2, 1992	61	9
TCE, Layer 1, 1986	143	29
TCE, Layer 2, 1986	92	13
DCE, Layer 1, 1986	108	23
DCE, Layer 2, 1986	59	7

$$Z = \frac{1}{2} \ln \left[\frac{1 + r}{1 - r} \right] \tag{8}$$

The sampling distribution of Z is approximated by (Hines and Montgomery 1980, p. 384; Guttman et al. 1982, p. 368)

$$Z \sim \text{Normal} \left(\frac{1}{2} \ln \left[\frac{1 + \rho}{1 - \rho} \right], \frac{1}{N - 3} \right) \tag{9}$$

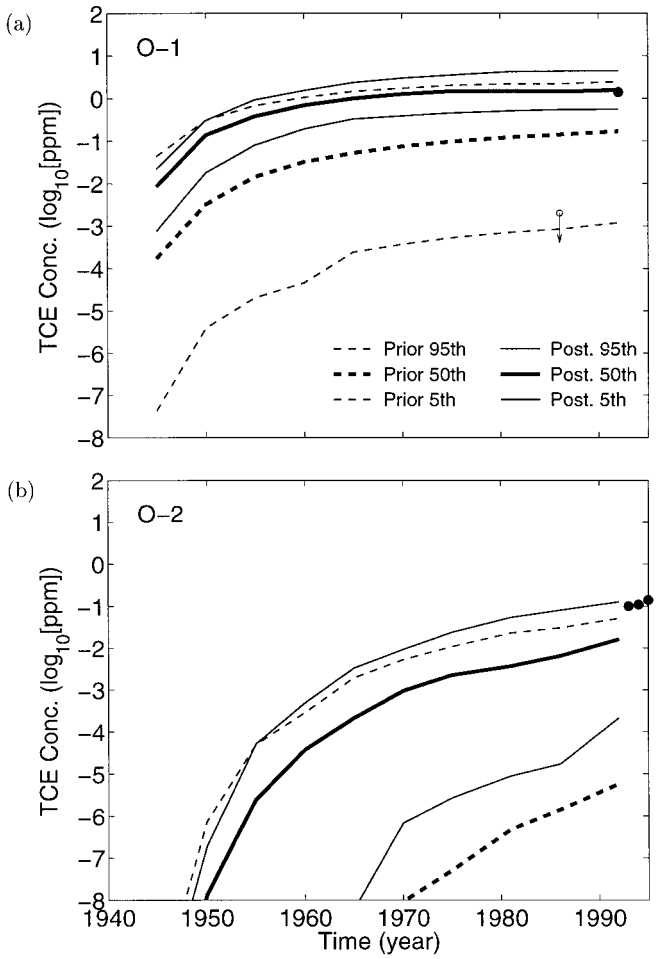

FIG. 5. Time Series Plots of Prior (Dotted Lines) and Posterior (Solid Lines) TCE Predictions of Median (50th Percentile of Uncertainty Distribution) and 5th to 95th Percentile Credible Intervals at Observation Points O-1 (a) and O-2 (b) (see Fig. 1); 5th Percentile Credible Limit of Prior at O-2 (b) Is Not Plotted on Log-Scale Axis Because Concentration Values Are Very Close to Zero; Solid Circles Represent Measurements; Downward-Pointing Arrow Represents Nondetection With Open Circle Tail at Detection Limit

where ρ = correlation coefficient between the predicted and observed (ln) concentrations expected when the model fit is ''perfect,'' and only measurement errors (in the measurements themselves) and errors due to the unmodeled spatial variability within a grid cell contribute to differences between the model and the data. This ''perfect model'' correlation coefficient is computed as

$$\rho = \left[\frac{\sigma^2_{c_o} - \sigma^2_{\varepsilon}}{\sigma^2_{c_o}}\right]^{1/2} \quad (10)$$

where $\sigma^2_{c_o}$ = variance of the measured concentrations across the site.

Eqs. (7) and (9) include the sample size, *N*, in the calculation of their respective sampling distributions and variances. However, correlations in the observed data, because of the spatial pattern and slow movement of the concentration field and the highly clustered sampling of the field, cause the sampling distribution variances to be greater than the values computed by these equations. This increase in variance can be captured by replacing the actual sample size, *N*, with a smaller effective sample size, N_{eff}. The effective sample size reflects the equivalent amount of independent information contained in the dataset. It is calculated based on the increase in the variance of the mean computed to occur when the concentration field is correlated compared with the case when the measurements are independent (Rodríguez-Iturbe and Mejía 1974; Bras and Rodríguez-Iturbe 1985). Details of this calculation are found in Sohn (1998). Table 3 summarizes the resulting effective sample sizes computed for the data used in the updating. The effective sample sizes are much lower than the actual number of samples in each case, because of the high degree of correlation induced by a highly structured sampling of a highly clustered concentration field (Fig. 3).

Results

The calculated mean value of the 1,200 prior predictions of TCE concentration in Layer 1 in 1992 is shown in Fig. 4(a). A posterior TCE plume is computed as the mean value of the corresponding updated (i.e., posterior) predictions and is shown in Fig. 4(b). Comparison of these two figures and the measured TCE concentrations (Fig. 3) shows that the much more elongated plume predicted as a result of the update is in better agreement with the observed data. The shape and mean concentrations of the prior and posterior TCE plumes in Layer 2 are similar to those in Layer 1, so are not illustrated here. The DCE predictions are also similar to those for TCE. Fig. 5 illustrates time series plots of TCE that summarize the historical reconstruction of the TCE contamination and its uncertainty at two locations along the major axis of the plume (the locations are identified in Fig. 1). The median posterior results of the model show better agreement with data at these locations than at the priors. The uncertainty is also significantly reduced, as reflected in the reduced width of the upper and lower bounds of the posterior predictions.

Table 2 summarizes the results of the source scenario uncertainty update. While a full discrimination among the scenarios is not achieved, the updating does favor Scenario 3 (TCE source in Layers 1 and 2 and no DCE source) as the most likely of the release scenarios, followed by Scenario 1 (TCE in Layer 1 only, no DCE source). As such, the updating favors a release scenario without DCE as a primary source. While scenarios with no DCE source (Scenarios 1 and 3) have 50% prior probability, their posterior probability increases to 71% (24.9 + 45.7%). This is also consistent with the fact that there are no recorded operations at the site that used significant quantities of DCE. Additionally, the measured DCE was in most cases solely the *cis* isomer, which is favored over the *trans* isomer during TCE biodegradation (Gossett and Zinder 1997; Carr and Hughes 1998). This suggests that the observed DCE is most likely present as a reaction byproduct of TCE degradation.

Similarly, the updating tends to favor a source of TCE in both the first and second aquifers. While the scenarios with TCE sources in Layers 1 and 2 (Scenarios 3 and 4) have a prior probability of 50%, their posterior probability increases to 61% (45.7 + 15.4%). This is consistent with the observation that at several locations in the vicinity of SA-1 and SA-2, the concentrations of TCE are higher in the second aquifer than in the first aquifer. Also, TCE concentrations greater than 10% of its aqueous solubility, an indication of DNAPL occurrence, are consistently detected over extensive areas in both the upper and lower aquifers. Hence, the preference for scenarios with sources in both aquifers—i.e., with DNAPL present in both aquifers—is consistent with the field evidence.

Table 4 contrasts the prior and posterior mean and standard deviation of input model parameters. The shift in the mean and the reduction in the parameter uncertainties suggest a posterior plume that moves faster than does the prior. For example, the median value of K_d decreases from approximately 0.75 l/kg to less than 0.5 l/kg, with a corresponding reduction in uncertainty. However, in some parameters little change in the parameter median or uncertainty is observed (e.g., see porosity

TABLE 4. Summary of Prior and Posterior Parameter Mean and Uncertainty

Parameter (1)	Prior		Posterior		% Change in standard deviation (6)
	Mean (2)	Standard deviation (3)	Mean (4)	Standard deviation (5)	
Longitudinal dispersivity (=α_l) (m)	49.7	27.5	64.9	16.4	−40
Transverse/longitudinal dispersivity ratio	0.25	0.25	0.22	0.22	−12
Vertical/longitudinal dispersivity ratio	0.028	0.026	0.023	0.019	−27
Sorption coefficient (=K_d) (l/kg)	0.984	0.794	0.432	0.162	−80
Porosity	0.30	0.028	0.29	0.026	−7
Near source TCE half-life (years)	12.1	13.1	15.0	12.1	−8
Far source TCE half-life (years)	27.4	14.1	22.6	10.7	−24
Near source DCE half-life (years)	10.6	12.5	22.2	13.4	+7
Far source DCE half-life (years)	24.4	13.9	31.5	10.4	−25
Source area concentrations:					
Area 1 TCE initial concentration (ppm)	272	212	154	96	−55
Area 2 TCE initial concentration (ppm)	18	10.3	28	10.7	+4
Area 3 TCE initial concentration (ppm)	0.033	0.026	0.046	0.021	−9
Area 1 DCE initial concentration (ppm)[a]	1,516	1,314	1,131	1,234	−6
Area 2 DCE initial concentration (ppm)[a]	17.2	10	17.7	10.3	+3
Area 3 DCE initial concentration (ppm)[a]	0.032	0.026	0.025	0.021	−19

[a]Applies only to Scenarios 2 and 4, where DCE sources are present (see Table 2).

TABLE 5. Summary of Prior and Posterior (i.e., Before and After Chemical Data Update) Probabilities for Hydraulic Conductivity Field Realizations

Hydraulic conductivity fields (1)	Geological structure (2)	Prior (3)	Posterior (4)
1–3	Anisotropic	30%	77.8%
4–8	Isotropic	50%	21.9%
9	Anisotropic-mean	10%	0.3%
10	Anisotropic-unlikely	10%	<0.00%

Note: Hydraulic conductivity fields 1–8 are most likely fields from head updating. Field 9 is mean hydraulic conductivity field from prior 1,000 anisotropic hydraulic conductivity fields. Field 10 is most unlikely field from head updating.

in Table 4), suggesting little sensitivity of the model predictions (relative to the observed data) for that parameter.

Table 5 summarizes the prior and posterior probabilities for the 10 hydraulic conductivity fields evaluated in the updating procedure. Significant results from this update are as follows:

1. The most unlikely hydraulic conductivity field (Realization #10) has the smallest posterior probability (effectively zero probability). Hence, the chemical updating is able to discriminate between reasonable (Realizations #1–9) and unreasonable hydraulic conductivity fields.
2. The mean anisotropic hydraulic conductivity field (Realization #9) has a very low posterior probability. The mean anisotropic hydraulic conductivity field represents a ''best guess'' hydraulic conductivity field from the hydraulic conductivity data (i.e., using ordinary kriging). However, the updating does not select this field as a highly likely hydraulic conductivity field for the chemical transport. One explanation for this is that the mean hydraulic conductivity field has smooth hydraulic conductivity transitions between adjacent cells, which cannot capture the abrupt changes in hydraulic conductivity and the associated increase in macroscale dispersion that occur in a heterogeneous hydraulic conductivity field. This suggests that if a mean hydraulic conductivity field is used, an approach that accounts for additional field-scale dispersivity is needed (see, e.g., Kitanidis and Vomvoris 1983).
3. The three anisotropic hydraulic conductivity fields (Realizations #1–3) have a combined posterior probability of 77.8% (30% prior), while the five isotropic hydraulic conductivity fields (Realizations #4–8) have a combined posterior probability of 21.9% (50% prior). Updating using chemical measurements thus favors the anisotropic hydraulic conductivity fields as the more likely geological structure, which is supported by regional scale studies conducted at the site (PRC 1992).

Model Projections

As an example of a typical application for the results of the updated model predictions, 30-year transport projections were made given current conditions and assuming that a local remediation activity was instituted to contain the sources in Areas 2 and 3 (migration from sources in the upper and lower aquifers in Area 1 were contained in 1987). From the 1,200 chemical transport simulations, 42 simulations made up 95% of the posterior probability. Only these 42 simulations were evaluated for the model projections. The updating procedure thus facilitated a significant reduction in the number of simulations needed to evaluate multiyear, multiscenario chemical transport projections, without losing much resolution in either the mean and/or the uncertainty in the projections. Tests were

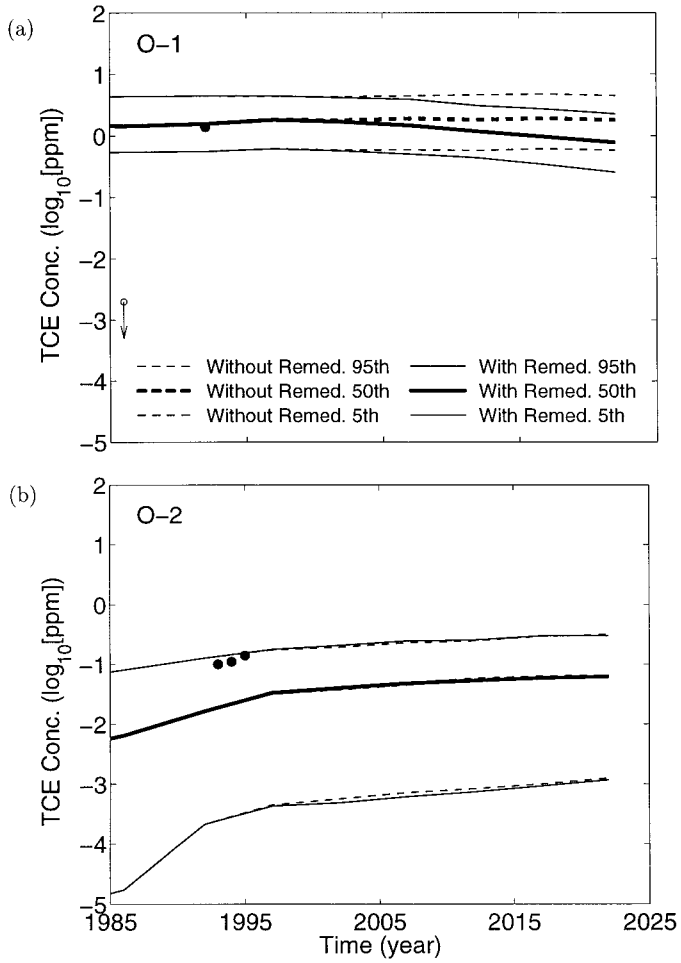

FIG. 6. Time Series Plots of Posterior TCE Projections Without (Dotted Lines) and With (Solid Lines) Aggressive Source Containment at O-1 (a) and O-2 (b) (see Fig. 1); Thick Lines Represent Median (50th Percentile of Uncertainty Distribution) Concentration and Thin Lines Represent 5th to 95th Percentile Credible Intervals; Solid Circles Represent Measurements; Downward-Pointing Arrow Represents Nondetection With Open Circle Tail at Detection Limit

conducted to ensure that omitting the simulations associated with the remaining 5% of the posterior probability did not significantly alter the model projections (see Sohn 1998). Fig. 6 illustrates time series plots of model predictions at O-1 and O-2 (the locations are identified in Fig. 1). The modeled source containment in Areas 2 and 3 results in notable reductions in ground-water concentrations only at O-1, the observation point closest to the contained source; yet even here the predicted median concentrations are within the uncertainty bounds for the no-containment case. These results are illustrative of the types of insights that can be achieved regarding remediation options, their uncertainty, and the need and potential value of further uncertainty reductions.

CONCLUDING REMARKS

This work investigated the application of Bayes Monte Carlo methods to complement site characterization and remedial investigation studies. The paper demonstrates methods for appropriately incorporating available data in Bayesian updating and indicates the type and degree of uncertainty reduction that can be achieved by following these procedures at a real hazardous waste site.

Appropriate updating procedures vary depending on the intended use of the updating and possible data correlations. A two-stage flow and transport modeling and updating procedure

was followed in this application. Flow predictions were updated using average observed hydraulic head gradients to achieve an initial reduction in the uncertainty of the hydraulic conductivity structure. Transport predictions were updated using chemical concentration data and their statistics selected so as to minimize the effects of temporal and spatial correlations. For this purpose, an approach was introduced that accounts for spatial correlation in the observed data by reducing the effective sample size.

The chemical data updating resulted in uncertainty reductions in most transport and reaction rate parameters, with reductions in standard deviations for all but three of the parameters, ranging 6 to 80%. Significant increases in posterior model probabilities associated with selected scenarios reduced the uncertainty in the considered alternatives as well. For example, the probability for the anisotropic hydraulic conductivity structure increased from a prior of 30% to a posterior of 78%. Similarly, the hypothesized presence of DCE solely as a by-product of TCE degradation was further supported by an increase in the probability of TCE-only source models from 50 to 71%.

The model projections with and without further remediation demonstrate a natural application of the results of the updating. Since 95% of the posterior uncertainty could be characterized by evaluating the 42 most significant realizations, out of a total prior set of 1,200, the updating procedure allowed for a significant reduction in the computational burden needed to evaluate multiyear, multiscenario chemical transport projections.

Possible candidate applications for this approach include (1) sites with long-term monitoring plans where successive updating would improve model predictions and uncertainty estimates; (2) sites where extensive field data are available to improve Monte Carlo simulation predictions; and (3) studies that require comprehensive site characterization where Bayes updating can be used to identify further data collection that can most effectively reduce overall uncertainty.

ACKNOWLEDGMENTS

This research was supported in part by NSF grant CMS-9502546. Computer support was provided in part by a grant from the Pittsburgh Supercomputing Center through the National Science Foundation (NSF) under grant BCS970001P. Development of the piezometric head fields benefited from discussions with Mark Cromer and from his modeling of the regional aquifer system. William Eddy and Ashish Sanil offered valuable advice for data analysis and model simulations. Simulation of the hydraulic conductivity fields was a collaborative effort with Grant Bromhal. We would like to thank Montgomery Watson and the U.S. Navy for providing the site data used in this study.

APPENDIX I. REFERENCES

Brand, K. P., and Small, M. J. (1995). ''Updating uncertainty in an integrated risk assessment: Conceptual framework and methods.'' *Risk Analysis*, 15(6), 719–731.

Bras, R. L., and Rodríguez-Iturbe, I. (1985). *Random functions and hydrology*, Addison-Wesley, Reading, Mass.

Carr, C. S., and Hughes, J. B. (1998). ''Enrichment of high-rate PCE dechlorination and comparative study of lactate, methanol, and hydrogen as electron donors to sustain activity.'' *Envir. Sci. Technol.*, 32(12), 1817–1824.

Chapra, S. C., and Reckhow, K. H. (1983). *Engineering approaches for lake management Volume 2: Mechanistic modeling*, Butterworth's, London.

Deutsch, C., and Journel, A. (1992). *GSLIB: Geostatistical software library and user's guide*, Oxford University Press, New York.

Essaid, H. I., and Hess, K. M. (1993). ''Monte Carlo simulations of multiphase flow incorporating spatial variability of hydraulic properties.'' *Groundwater*, 31(1), 123–143.

Franks, S. W., Gineste, P., Beven, K. J., and Merot, P. (1998). ''On constraining the predictions of a distributed model: The incorporation of fuzzy estimates of saturated areas into the calibration process.'' *Water Resour. Res.*, 34(4), 787–797.

Gossett, J. M., and Zinder, S. H. (1997). ''Microbiological aspects relevant to natural attenuation of chlorinated ethenes.'' *Proc. Symp. on Natural Attenuation of Chlorinated Organics in Ground Water: EPA/540/R-97/504*, U.S. Environmental Protection Agency, Washington, D.C., 12–15.

Guttman, I., Wilks, S. S., and Hunter, J. S. (1982). *Introductory engineering statistics*, 3rd Ed., Wiley, New York.

Hines, W. W., and Montgomery, D. C. (1980). *Probability and statistics in engineering and management science*, 2nd Ed., Wiley, New York.

Howard, P. H., Boethling, R. S., Jarvis, W. F., Meylan, W. M., and Michalenko, E. M. (1991). *Handbook of environmental degradation rates*, Lewis, Boca Raton, Fla.

James, B. R., and Freeze, R. A. (1993). ''The worth of data in predicting aquitard continuity in hydrogeological design.'' *Water Resour. Res.*, 29(7), 2049–2065.

Kitanidis, P. K., and Vomvoris, E. G. (1983). ''A geostatistical approach to the inverse problem in groundwater modeling (steady state) and one-dimensional simulations.'' *Water Resour. Res.*, 19(3), 677–690.

Loaciga, H. A., and Mariño, M. A. (1987). ''Parameter estimation in groundwater: Classical, Bayesian, and deterministic assumptions and their impact on management policies.'' *Water Resour. Res.*, 23(6), 1027–1035.

Marryott, R. A. (1996). ''Optimal ground-water remediation design using multiple control technologies.'' *Groundwater*, 34(3), 425–433.

Massmann, J., Freeze, R. A., Smith, L., Sperling, T., and James, B. (1991). ''Hydrogeological decision analysis: 2. Applications to groundwater contamination.'' *Groundwater*, 29(4), 536–548.

McDonald, M. G., and Harbaugh, A. W. (1988). *A modular three-dimensional finite-difference ground-water flow model. Techniques of water resources investigations, Book 6*, U.S. Geological Survey, Washington, D.C.

Montgomery, J. H. (1992). *Groundwater chemicals desk reference*, 2nd Ed., Lewis, Boca Raton, Fla.

Nyer, E. K., and Duffin, M. E. (1997). ''The state of the art of bioremediation.'' *GWMR*, 17(2), 64–69.

PRC. (1992). *PRC Environmental Management and James M. Montgomery, Consulting Engineers, Tech. Memo.: Geology and hydrogeology final draft*, PRC Environmental Management, Walnut Creek, Calif.

Reichard, E. G., and Evans, J. S. (1989). ''Assessing the value of hydrogeologic information for risk-based remedial action decisions.'' *Water Resour. Res.*, 25(7), 1451–1460.

Rodríguez-Iturbe, I., and Mejía, J. M. (1974). ''The design of rainfall networks in time and space.'' *Water Resour. Res.*, 10(4), 713–728.

Sohn, M. D. (1998). ''Updating uncertainty in site characterization and chemical transport using Bayesian Monte Carlo methods.'' PhD thesis, Carnegie Mellon University, Pittsburgh.

U.S. Environmental Protection Agency (USEPA). (1989). ''In-situ aquifer restoration of chlorinated aliphatics by methanotrophic bacteria.'' *EPA/600/2-89/033*, Washington, D.C.

Weisberg, S. (1985). *Applied linear regression*, 2nd Ed., Wiley, New York.

Yeh, W. W. (1986). ''Review of parameter identification procedures in groundwater hydrology: The inverse problem.'' *Water Resour. Res.*, 22(2), 95–108.

Yevjevich, V. (1972). *Probability and statistics in hydrology*, Water Resources Publications, Fort Collins, Colo.

Zheng, C. (1992). *A modular three-dimensional transport model: Documentation and user's guide*, S. S. Papadopoulos & Assoc., Inc., Bethesda, Md.

Zheng, C., Wang, P. P., and Dortch, M. (1997). ''Three-dimensional modeling of multicomponent transport in heterogeneous aquifers: The ULTIMATE conservative difference scheme.'' *Eos Trans. AGU Spring Meet. Suppl.*, 78(17), S169.

APPENDIX II. NOTATION

The following symbols are used in this paper:

$\bar{c}_m$ = mean of model predictions;
$\bar{c}_o$ = mean of observed (ln) concentrations;
K_d = sorption coefficient;
$L(\ )$ = likelihood;
N = sample size;
O = observation;
$p(Y_k)$ = prior probability mass for model output Y_k;
p'_u = posterior probability mass for realization u;
V, Y = model output or input parameter;
α = dispersivity;
μ' = posterior mean;
ρ, r = correlation coefficient;
ρ' = posterior correlation coefficient;
σ^2 = variance of measurement error; and
σ'^2 = posterior variance.